

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 11:14:25 ; Search time 6119 Seconds
(without alignments)
10783.989 Million cell updates/sec

Title: US-09-647-841B-1
Perfect score: 1613
Sequence: 1 ttttataaatatttaagctt.....tgaacttgagtacctaactc 1613

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1428.6	88.6	3617	8	AF028841	AF028841 Medicago
2	175.4	10.9	3234	8	PSPTXAG	X67427 Pisum sativo
3	117.8	7.3	95620	2	BX088600	BX088600 Danio rer
4	115.6	7.2	128769	2	AL139261	AL139261 Homo sapi
c 5	114.6	7.1	129240	9	AC084128	AC084128 Homo sapi
c 6	114	7.1	8056	6	AX599046	AX599046 Sequence
c 7	113.2	7.0	123589	9	AC104790	AC104790 Homo sapi
c 8	113.2	7.0	172758	2	AC022553	AC022553 Homo sapi
c 9	112.8	7.0	159948	2	BX248240	BX248240 Danio rer
10	112.8	7.0	187310	2	BX005122	BX005122 Danio rer
c 11	112.2	7.0	178066	2	BX004966	BX004966 Danio rer
12	111.2	6.9	112917	8	AC123976	AC123976 Medicago
c 13	110.8	6.9	196490	2	AC005507	AC005507 Plasmodiu
c 14	110.6	6.9	147956	2	AC137839	AC137839 Medicago
15	109.8	6.8	195032	2	AC144827	AC144827 Danio rer
16	109	6.8	169546	2	AC004157	AC004157 Plasmodiu
c 17	109	6.8	250707	3	AE014848	AE014848 Plasmodiu
c 18	108.8	6.7	180919	2	BX001038	BX001038 Danio rer
c 19	108.2	6.7	196859	9	AC087277	AC087277 Homo sapi
c 20	107.4	6.7	80150	9	AC134620	AC134620 Homo sapi
c 21	107.2	6.6	26561	9	AL929515	AL929515 Human DNA
c 22	106.8	6.6	181792	9	AC098822	AC098822 Homo sapi
c 23	106.6	6.6	136713	9	AC112255	AC112255 Homo sapi
c 24	106.4	6.6	241454	2	BX470172	BX470172 Danio rer
25	105.8	6.6	151815	2	BX470147	BX470147 Danio rer
26	105.6	6.5	101884	8	AC142224	AC142224 Medicago
c 27	105.6	6.5	188485	2	BX470246	BX470246 Danio rer
c 28	105.6	6.5	188716	5	AL807739	AL807739 Zebrafish
c 29	105.4	6.5	198583	2	AC144702	AC144702 Danio rer
c 30	105.2	6.5	60529	2	AC144825	AC144825 Danio rer
c 31	104.8	6.5	89728	5	AL672079	AL672079 Zebrafish
32	104.8	6.5	126175	2	AC136841	AC136841 Medicago
33	104.6	6.5	201591	2	AL954657	AL954657 Danio rer
34	104.6	6.5	313050	3	PFA929352	AL929352 Plasmodiu
c 35	104.4	6.5	187197	2	BX530723	BX530723 Danio rer
c 36	104.2	6.5	129240	9	AC084128	AC084128 Homo sapi
c 37	104	6.4	343050	3	PFA929353	AL929353 Plasmodiu
c 38	103.6	6.4	67970	3	PFMAL1P3	AL031746 Plasmodiu
c 39	103.6	6.4	250078	3	AE014829	AE014829 Plasmodiu
40	103.4	6.4	1867	8	MTSCAJ23	AJ223323 Saccharom
c 41	103.4	6.4	2144	5	BC044278	BC044278 Xenopus l
c 42	103.4	6.4	146039	5	AL627132	AL627132 Zebrafish
c 43	103.2	6.4	130110	9	AL731777	AL731777 Human DNA
c 44	103.2	6.4	152369	2	CNS07ED2	AL454443 Homo sapi
45	103.2	6.4	172816	9	AC093899	AC093899 Homo sapi

ALIGNMENTS

RESULT 1
AF028841
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AF028841
Medicago sativa
AF028841 L37017
AF028841.1 GI:3818415
Medicago sativa
Medicago sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1224 to 3617)
Deutch,C.E. and Winicov,I.

3617 bp DNA linear PLN 09-JAN-1999
Medicago sativa proline-rich cell wall protein gene, complete cds.

340	TTAAGCATATAACATCAAAAGTCTAAACATAGCAAAATGTTGTTTTTAGATGACACATT	399
889	TCAT--ATAGTTTAAAGGATGATGATTCGATTACAAAACAAAATACTAATAATCTAG	946
400	TCATACATAGTTTAAAGATACCTTGATTCGATTACAAAAGAAATTTACCAATAGTT-TAG	458
947	CACAAAGTTTAAAGCAATATATATAAGAGCTTCATAGCATGTGGATATTCATTTAGAATAT	1006
459	CACAAAGTCTAAAGCATAATTA--GATCAGATGTGCAGATT--TATGAAAAAA	511
1007	AGATTAGATTGCCCTTTTCATCAGCGGTC--TAACAGCACCACCTTGCTCACTACATGT--	1061
512	GATTAGATTGCCCTTTTCATCAGCGGTGAATAGCAGCACTACTTGTCACTACATGTTA	571
1062	CAAAAATGTCCTTAGTACAGACCGGCTTTTACTTTGATTCCTCCCTTGTCCATGCATGAAA	1121
572	AAAAATGTCCTTAGTACATCAAACTTTTTCATTTGATTCCTCCCTT--ATCCATGAAA	627
1122	AAAAATGTCCTTAGTACATCAAACTTTTTCATTTGATTCCTCCCTTGTCCATGCATGAAA	1181
628	AAAAATGTCCTTAGTACATCAAACTTTTTCATTTGATTCCTCCCTTGTCCATGCATGAAA	687
1182	TAGTTTG 1188	
688	TTGTTTG 694	
RESULT 3		
LOCUS	BX088600	
DEFINITION	Danio rerio clone DKEY-166M7, *** SEQUENCING IN PROGRESS ***, 5	
ACCESSION	BX088600	
VERSION	GI:30525002	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.	
SOURCE	Danio rerio (zebrafish)	
ORGANISM	Danio rerio	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.	
AUTHORS	McLaren,S.	
TITLE	Direct Submission	
JOURNAL	Submitted (29-MAY-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk	
COMMENT	On May 11, 2003 this sequence version replaced gi:2950041.	
	----- Genome Center	
	Center: Wellcome Trust Sanger Institute	
	Center code: SC	
	Web site: http://www.sanger.ac.uk	
	Contact: zfish-help@sanger.ac.uk	
	----- Project Information	
	Center project name: zK166M7	
	----- Summary Statistics	
	Assembly program: XGAP4; version 4.5	
	Chemistry: Dye-terminator; 100% of reads	
	Consensus quality: 192251 bases at least Q40	
	Consensus quality: 192839 bases at least Q30	
	Consensus quality: 193366 bases at least Q20	
	Insert size: 195220; sum-of-contigs	
	Insert size: 189846; 4.6% error; agarose-fp	
	Quality coverage: 6.33x in Q20 bases; sum-of-contigs Quality	
	coverage: 6.55x in Q20 bases; agarose-fp	

	* NOTE: This is a 'working draft' sequence. It currently	
	* consists of 5 contigs. The true order of the pieces	
	* is not known and their order in this sequence record is	
	* arbitrary. Gaps between the contigs are represented as	
	* runs of N, but the exact sizes of the gaps are unknown.	
	* This record will be updated with the finished sequence	
	* as soon as it is available and the accession number will	
	* be preserved.	

QY 533 TATAATTACATGCCGTTACGGTAAAAAATGGATAAAATTGGGTATGGAGTACTAGTAATT 592

misc_

```
feature
1. .33392
/clone_lib="RPCI-5"
1. .33392
```

[illegible]


```

QY 961 CAATATTATAAGCTTCATAGCATGCGATATTCATTAGAAAATATAGATTGCCC 1020
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 902 TATAAATTTAAATATACAAATTTTATACATAAAAAAATAAAAAACAATTTT 843
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1021 CTTTCATCAGCGGTCTAACAGCACCACTGTCTACTACATGTCAAAAATGCTCTAGTAC 1080
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 842 TTTAAATAAAATATTTTAAATAATCATTTTATAAATTTTATTTAAATATAAATAAAT 783
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1081 AGCACCGCTTTTACTTGATTCCTCCCTGTCCATGCATGAAAAAATCAAAACAATATTG 1140
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 782 AATTTTATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 723
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1141 GACACACAACTTCCGCCCACTTCTCTTTCTCTGCTGCTAGTTGTTGAGACATCAT 1200
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 722 TAAATTAATTTTAAATATAAAAAAATTTTATATTTTATTTTATTTTATTTTATTTT 663
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1201 ATTGA-TCAATTTGGCTGATGAATTCACACAAAAAATTCACCTACCCATTCATGCTGTG 1259
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 ATAAATTTTATTTTAAATTTTATTTTAAATTTTAAATTTTAAATTTTATATTTTATTT 603
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1260 GGGCCACATATAAATCCATGAGGATTCAT-GTCCATCCCAAGTCAATGATTCACAT 1318
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 TAAATTTATTTAATATAAATTTTAAATTTTAAATTTTAAATTTTATATTTTATTTAAT 543
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1319 ATATAACATTTGAATTAATTTTAAATTTTCCAGTATTATGATTTAGATTGATT 1372
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 TTTAAATTTAATATAAATTTTAAATTTTAAATTTTAAATTTTATATTTTATTTAATTTT 489
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
AC104790/c
LOCUS Homo sapiens BAC clone RP11-167A8 from 4, complete sequence.
DEFINITION AC104790
ACCESSION AC104790.4 GI:20258589
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Sulston,J.E. and Waterston,R.
1 (bases 1 to 123589)
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 123589)
AUTHORS Pearman,C., Meyer,R. and Doeber,A.
TITLE The sequence of Homo sapiens BAC clone RP11-167A8
JOURNAL Unpublished (2001)
MEDLINE 3 (bases 1 to 123589)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 123589)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (21-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 123589)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Apr 21, 2002 this sequence version replaced gi:18652549.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc

```

Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0167A08

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Cataneese, J.J. and de Jong, P.J. (1998). An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-460L16; the clone sequenced to the right is RP11-162F1, 2000 bp overlap. Actual start of this clone is at base position 41945 of RP11-460L16; actual end is at base position 44749 of RP11-162F1.

Polymorphisms exist between AC104790 and AC104804. Data from AC104804 and AC106867 was used to finish AC104790.

FEATURES

	source
repeat_region	1. .123589
repeat_region	/organism="Homo sapiens"
repeat_region	/mol_type="genomic DNA"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="4"
repeat_region	/map="4"
repeat_region	/clone="RP11-167A8"
repeat_region	/clone_lib="RPC1-11"
repeat_region	353. .379
repeat_region	/rpt_family="AT-rich"
repeat_region	1639. .1719
repeat_region	/rpt_family="L1"
repeat_region	1760. .1953
repeat_region	/rpt_family="L1"
repeat_region	1992. .2398
repeat_region	/rpt_family="L1"
repeat_region	4331. .4798
repeat_region	/rpt_family="ERV1"
repeat_region	5422. .5963
repeat_region	/rpt_family="ERV1"
repeat_region	5964. .6006
repeat_region	/rpt_family="(TG)n"
repeat_region	6007. .6703
repeat_region	/rpt_family="ERV1"
repeat_region	7219. .7695
repeat_region	/rpt_family="ERV1"
repeat_region	7860. .7909
repeat_region	/rpt_family="L2"
repeat_region	7910. .8380
repeat_region	/rpt_family="MALR"

[illegible]

RESULT 10

LOCUS	187310 bp	DNA	linear	HTG 02-JUN-2003
DEFINITION	Danto reri1o clone CH211-134D11, *** SEQUENCING IN PROGRESS ***, 4 unordered pieces.			

EX005122.7 GI:29500520
EX005122.7 GI:29500520
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 187310)
McLaren, S.
Direct Submission
Submitted (29-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Apr 2, 2003 this sequence version replaced gi:28272911.

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces


```

REFERENCE
AUTHORS      2 (bases 1 to 112917)
              Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
              Cook,D., Kim,D. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (06-JUN-2002) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
REFERENCE
AUTHORS      3 (bases 1 to 112917)
              Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
              Cook,D., Kim,D. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (12-NOV-2002) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
REFERENCE
AUTHORS      4 (bases 1 to 112917)
              Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
              Cook,D., Kim,D. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (08-MAR-2003) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
REFERENCE
AUTHORS      5 (bases 1 to 112917)
              Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
              Cook,D., Kim,D. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (10-JUN-2003) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
REFERENCE
AUTHORS      6 (bases 1 to 112917)
              Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
              Cook,D., Kim,D. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (11-JUN-2003) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
COMMENT      On Jun 11, 2003 this sequence version replaced gi:31560173.
              -----
              Center: Genome Center
              Department Of Chemistry And Biochemistry
              The University Of Oklahoma
              Center code:UOKNOR
              -----
FEATURES
source       1..112917
              /organism="Medicago truncatula"
              /mol_type="genomic DNA"
              /db_xref="taxon:3880"
              /clone="mth2-30123"
              /clone_lib="Medicago truncatula BAC library H2"
BASE COUNT   38546 a 17840 c 18400 g 38131 t
ORIGIN
Query Match      6.9%; Score 111.2; DB 8; Length 112917;
Best Local Similarity 68.5%; Pred. No. 3.3e-06;
Matches 152; Conservative 1; Mismatches 69; Indels 0; Gaps 0;
QY      235 TATAGTACATCAACACCTTTGAGAGGTTAAGTGTGCGGATTTCTAGATAACAAG 294
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      22909 TGTAGTGAACACACATCTTGTGAGAGGTAACAAGTGTGCAGATTCCTTAGTACGCCA 22968
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      295 GYGCATTCCAGATCTTCTTGTGCGAGCTTGAGAACCCCTATCCTGGCGCTTGGAGATT 354
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      22969 GTGCTATTGATGCTTCTTGTGCGAGCTTGTGAACCTCATCTTATGCTTGGAGATT 23028
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      355 TACTTCTTGTGATGCTTCTAGAGTACAGCTCCTTAAAGGCTGTAGTCTAGTCTATTTT 414
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      23029 TACTTATTGATGCTTCTAGAGTACAGCTTCTTAAAGGCTGTAGTCTATTTTCT 23088
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      415 CATCCTTCTACAAAAAAGTCAFAAATATAGTTTA 456
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      23089 TCCTGCACTAGTAATAAAGATAAATAGTAACAAAGATTA 23130
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 13

```

```

AC005507/c
LOCUS        196490 bp DNA linear HTG 12-AUG-2000
DEFINITION   Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
              PROGRESS **, 2 unordered pieces.
AC005507
VERSION      AC005507.8 GI:9797737
KEYWORDS     HTG; HTGS_PHASE1.
SOURCE       Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM     Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE    1 (bases 1 to 196490)
AUTHORS      Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T.,
              Kurdi,O.B., Conway,A.B. and Davis,R.W.
TITLE        Plasmodium falciparum 3D7 chromosome 12
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 196490)
AUTHORS      Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
TITLE        Direct Submission
JOURNAL      Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
              Center, Stanford University, 855 California Avenue, Palo Alto, CA
              94304, USA
COMMENT      On Aug 12, 2000 this sequence version replaced gi:8810456.
              * NOTE: This is a 'working draft' sequence. It currently
              * consists of 2 contigs. The true order of the pieces
              * is not known and their order in this sequence record is
              * arbitrary. Gaps between the contigs are represented as
              * runs of N, but the exact sizes of the gaps are unknown.
              * This record will be updated with the finished sequence
              * as soon as it is available and the accession number will
              * be preserved.
              * 1 181303: contig of 181303 bp in length
              * 181304 181503: gap of unknown length
              * 181504 196490: contig of 14987 bp in length.
              Location/Qualifiers
              1..196490
              /organism="Plasmodium falciparum"
              /mol_type="genomic DNA"
              /db_xref="taxon:5833"
              /chromosome="12"
              /clone="PFYACB8-628"
              /clone="3D7"
BASE COUNT   77103 a 20186 c 18680 g 80315 t 206 others
ORIGIN
Query Match      6.9%; Score 110.8; DB 2; Length 196490;
Best Local Similarity 64.5%; Pred. No. 3.2e-06;
Matches 182; Conservative 0; Mismatches 97; Indels 3; Gaps 1;
QY      2 TTTTAAATATTTTAACTTGTGTAATAAATTTTGGGATCTATATATA---AGCCCACTACCA 58
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      182646 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 182587
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      59 ATTTTAAATATATATATATATATATATATATATATATATATATATATATATATATAT 118
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      182586 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 182527
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      119 ATTTTAAATATATATATATATATATATATATATATATATATATATATATATATAT 178
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      182526 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 182467
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      179 TATTACGTTGATGGTAAAAAATAAATAAATTTGTTACCACTTTAAAGTCATAATATA 238
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      182466 ATTTTAAAGATTTGTACATAAACAATAACGTTCAATTAATTTTAAATCATTTATA 182407
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      239 GTACAATCCAAACCCCTTGGAGAGGTTAATGCTGTGCGGATTT 280
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      182406 ATATCTTACCTTCATGTTTAAATTTTAAATGTTTAAAGAATTT 182365
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 14
AC137839/c
LOCUS        147956 bp DNA linear HTG 27-MAY-2003
DEFINITION   Medicago truncatula clone mch2-34j11, WORKING DRAFT SEQUENCE, 18

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

unordered pieces.
AC137839
AC137839.10 GI:31076555
HTG: HTGS_PHASE1; HTGS_DRAFT.
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Medicago truncatula BAC Clone mth2-34j11
Unpublished
2 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (04-DEC-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (27-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 27, 2003 this sequence version replaced gi:30270629.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Medicago truncatula BAC Clone mth2-34j11
Unpublished
2 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (04-DEC-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (27-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 27, 2003 this sequence version replaced gi:30270629.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Medicago truncatula BAC Clone mth2-34j11
Unpublished
2 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (04-DEC-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (27-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 27, 2003 this sequence version replaced gi:30270629.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Medicago truncatula BAC Clone mth2-34j11
Unpublished
2 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (04-DEC-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (27-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 27, 2003 this sequence version replaced gi:30270629.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

COMMENT

----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2190: contig of 2190 bp in length
* 2191 2290: gap of unknown length
* 2291 5232: contig of 2942 bp in length
* 5233 5332: gap of unknown length
* 5333 7396: contig of 2064 bp in length
* 7397 7496: gap of unknown length
* 7497 10640: contig of 3144 bp in length
* 10641 13741: gap of unknown length
* 13741 13777: contig of 2637 bp in length
* 13778 13477: gap of unknown length
* 13478 16887: contig of 3210 bp in length
* 16888 16787: gap of unknown length
* 16788 19875: contig of 3088 bp in length
* 19876 19975: gap of unknown length
* 19976 23802: contig of 3827 bp in length
* 23803 23902: gap of unknown length
* 23903 27548: contig of 3646 bp in length
* 27549 27648: gap of unknown length
* 27649 33103: contig of 5455 bp in length
* 33104 33203: gap of unknown length
* 33204 41209: contig of 8006 bp in length
* 41210 41309: gap of unknown length
* 41310 47453: contig of 6144 bp in length
* 47454 47553: gap of unknown length
* 47554 57436: contig of 9883 bp in length
* 57437 57536: gap of unknown length
* 57537 67037: contig of 9501 bp in length
* 67038 67137: gap of unknown length
* 67138 78583: contig of 11446 bp in length
* 78584 78683: gap of unknown length

FEATURES
source

* 78684 95337: contig of 16654 bp in length
* 95338 95437: gap of unknown length
* 95438 121254: contig of 25817 bp in length
* 121255 121354: gap of unknown length
* 121355 147956: contig of 26602 bp in length.

Location/Qualifiers
1..147956
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-34j11"
/clone_lib="Medicago truncatula BAC library H2"
BASE COUNT 48249 a 24438 c 25475 g 48056 t 1738 others
ORIGIN

Query Match 6.9%; Score 110.6; DB 2; Length 147956;
Best Local Similarity 61.2%; Pred. No. 3.6e-06;
Matches 194; Conservative 1; Mismatches 120; Indels 2; Gaps 1;

QY 228 TCATAAATATAGTACAATCCCAACCCCTTTGAGAGGTTAATGTGTGCGGATTTCTAGAT 287
DB 111166 TGAATGTTCATTTTACCACACCATTCGAGAGACAAATGAGTGTAGACTTTTACGT 111107
QY 288 AAACAAGTGGCAATTCAGATTTCTTTGGTGCAGCTGGAGAACCCATCTCTGGCTTG 347
DB 111106 AAGCATGTGCGCAATCAGACTCTTTCATGGTGCAGTGGAGAACCCCTCTCCGGCTTG 111047
QY 348 GAAGATTTTACTTCTTGTGATGCTTCTAGAGTACAGCTCTTAAAGCTGTAGTCTAGTTT 407
DB 111046 GAAGCTTTTACTTCTTCTCCTCAATGTCATCTAGAGTGTAGTTTC--TTAGCGGCTAGTCTAAATTT 110989
QY 408 TTTTTCATCTCTCTACCAAAAAAAGTATATAATAGTTTATACATATAACT 467
DB 110988 TCTGTTTTCCTTCTCTTCTTCTTATTTACCAAAAAACAAAAAGATATATAA 110929
QY 468 TTAATAAATAAATAAATAAATTTTCATCCCTAAACACATAGTAGATAAATTTTCAAAAAAATA 527
DB 110928 TCAGTTTCTTAAATAAATAAATACAGCATGMAAACTCTGAAAAACCAACACACATTTTC 110869
QY 528 TTGTTTATAATTTACAT 544
DB 110868 CTATCTCAATTTGGATAT 110852

RESULT 15

AC144827

LOCUS

DEFINITION

AC144827

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

195032 bp DNA linear HTG 22-MAY-2003
Danio rerio clone CH211-77C11, *** SEQUENCING IN PROGRESS ***, 7
ordered pieces.
AC144827
HTG: HTGS_PHASE2; HTGS_ACTIVEFIN.
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 195032)
Talbot,W.S., Rauch,G.J., Grimwood,J., Dickson,M., Schmutz,J. and
Myers,R.M.
Genomic studies of vertebrate diversity
Unpublished
2 (bases 1 to 195032)
Talbot,W.S., Rauch,G.J., Grimwood,J., Dickson,M., Schmutz,J. and
Myers,R.M.
Direct Submission
Submitted (22-MAY-2003) Stanford Human Genome Center, 975
California Avenue, Palo Alto, CA 94304, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have

```

* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 61516: contig of 61516 bp in length
*
* 61517 61616: gap of unknown length
*
* 61617 117922: contig of 53306 bp in length
*
* 117923 118022: gap of unknown length
*
* 118023 125117: contig of 7095 bp in length
*
* 125118 125217: gap of unknown length
*
* 125218 126859: contig of 1642 bp in length
*
* 126860 126959: gap of unknown length
*
* 126960 134961: contig of 8002 bp in length
*
* 134962 135061: gap of unknown length
*
* 135062 147553: contig of 12492 bp in length
*
* 147554 147653: gap of unknown length
*
* 147654 195032: contig of 47379 bp in length.
The sequence of the clone was established as a mapping and
sequencing collaboration at the Stanford Genome Evolution Center,
funded by the NIH Centers of Excellence in Genomic Science (CEGS)
Initiative (http://cegs.stanford.edu). The clone was isolated from
the BAC library CHORI-211 (http://bacpac.chori.org).
-----Genome Center
Center: Stanford Human Genome Center
Center Code: shgc
Web site: http://www-shgc.stanford.edu
-----
Project Information
Center Project Name: 2341
Center clone name: CH211-77C11
-----
Summary Statistics
Consensus quality: 191969 bases at least Q40
Consensus quality: 192317 bases at least Q30
Consensus quality: 192756 bases at least Q20
Estimated insert size: 190000; agarose-fp estimation
Estimated insert size: 193122; sum-of-contigs
estimation
Quality coverage: 1.01 in Q20 bases; agarose-fp
estimation
Quality coverage: 1 in Q20 bases; sum-of-contigs
estimation.
Location/Qualifiers
1. 195032
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-77C11"
BASE COUNT 61347 a 35379 c 36159 g 61547 t 600 others
ORIGIN
Query Match 6.8%; Score 109.8; DB 2; Length 195032;
Best Local Similarity 42.9%; Pred. No. 4.3e-06;
Matches 412; Conservative 0; Mismatches 542; Indels 7; Gaps 3;
QY 7 AAATATTTAAGCTTGATATAAATTTTGCATCTATATATAAGCCACATACCAATTTAAAA 66
Db 134453 AAACAATAATGTGTCACATTAATAATGAATTTCTATAGTATATCAACAATATATAT 134512
QY 67 TTATATATATATATATATATATATATATATATAATAATTTTATTTACCAATTTAAAA 126
Db 134513 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 134572
QY 127 TTATATATATATATATATATATATATATATATAATAATTTTAT-TATATTTATTAG 185
Db 134573 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 134632
QY 186 TTGATGGTAAAAAATAAATATAATTTGTTACCATTTAAAAAGTCATAAATATAGTACAA 245
Db 134633 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 134692
QY 246 CCAACCCCTTTGAGAGGTTAATGTGTGCGGATTTCTAGATAAACAAGGYCCCAATTCAC 305

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 10:52:01 ; Search time 482 Seconds
(without alignments)
9033.597 Million cell updates/sec

Title: US-09-647-841B-1

Perfect score: 1613

Sequence: 1 ttttataatatttaagctt.....tgaacttgagtaacttactt 1613

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 252756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1427	88.5	1612	21	AAZ34539
c	114	7.1	8056	25	ABZ10246
c	100.6	6.2	16258	24	ABL70376
c	100.6	6.2	16258	24	ABK40038
5	100.2	6.2	8056	25	ABZ10246
c	95.2	5.9	8056	25	ABZ10100
7	93.2	5.8	6641	24	ABL54336
8	93.2	5.8	6641	24	ABL32315

c	9	92.4	5.7	9770	24	ABL32032	Human immune syste
c	10	92	5.7	6621	24	ABL70156	Chemically treated
c	11	92	5.7	6621	24	ABK33933	Human DNA for stag
c	12	92	5.7	6621	24	AAS61098	Human gene regulat
c	13	92	5.7	6621	25	ABZ09984	Haematopoietic cel
c	14	92	5.7	6621	25	ABZ10098	Haematopoietic cel
c	15	92	5.7	6621	25	ABZ10130	Haematopoietic cel
c	16	92	5.7	6621	25	ABZ10244	Haematopoietic cel
c	17	91.6	5.7	11745	24	ABK28332	DNA transcription
c	18	91.4	5.7	6641	24	ABL54336	Chemically treated
c	19	91.4	5.7	6641	24	ABL32315	Human immune syste
c	20	91	5.6	3991	22	AAD16633	Human novel protei
c	21	90.6	5.6	8056	25	ABZ10100	Haematopoietic cel
c	22	90.6	5.6	8136	24	ABK39957	Human chemically p
c	23	90.6	5.6	8136	24	ABL32555	Human immune syste
c	24	90	5.6	11745	24	ABK28332	DNA transcription
c	25	90	5.6	18060	24	ABL92212	Chemically treated
c	26	90	5.6	18060	24	AAD22313	Chemically treated
c	27	89.4	5.5	6621	24	ABL70156	Human DNA for stag
c	28	89.4	5.5	6621	24	ABK33933	Human gene regulat
c	29	89.4	5.5	6621	24	AAS61098	Haematopoietic cel
c	30	89.4	5.5	6621	25	ABZ09984	Haematopoietic cel
c	31	89.4	5.5	6621	25	ABZ10098	Haematopoietic cel
c	32	89.4	5.5	6621	25	ABZ10130	Haematopoietic cel
c	33	89.4	5.5	6621	25	ABZ10244	Haematopoietic cel
c	34	89.4	5.5	113445	22	AAI61373	Soybean 318013 reg
c	35	89.2	5.5	8136	24	ABK39957	Human chemically p
c	36	89.2	5.5	8136	24	ABL32555	Human immune syste
c	37	88.2	5.5	3296	23	ABL17264	Drosophila melanog
c	38	87.4	5.4	6065	24	ABL32505	Human immune syste
c	39	87.4	5.4	14551	24	ABL34585	Human metastasis a
c	40	87.2	5.4	8136	24	ABK39956	Human chemically p
c	41	87.2	5.4	8136	24	ABL32554	Human immune syste
c	42	87.2	5.4	14551	24	ABL34585	Human metastasis a
c	43	87.2	5.4	18855	24	ABL32611	Human immune syste
c	44	86.8	5.4	169739	24	ABQ88186	Human osteoblast d
c	45	86.2	5.3	20420	22	AAK73165	Human immune/haema

ALIGNMENTS

RESULT 1
AAZ34539
ID AAZ34539 standard: DNA; 1612 bp.
XX AAZ34539;
AC AAZ34539;
XX 01-FEB-2000 (first entry)
XX Alfalfa salt inducible MsPRP2 gene promoter region.
XX MsPRP2 gene; promoter; Alfinl1; transcription factor; alfalfa;
XX salt tolerance; stress tolerance; transgenic plant; root; ds.
XX Medicago sativa.
XX Key Location/Qualifiers
XX protein_bind complement (718..727)
FT /tag= a
FT /note= "Alfinl binding site"
FT complement (778..786)
FT /tag= b
FT /note= "Alfinl binding site"
FT complement (1034..1049)
FT /tag= c
FT /note= "Alfinl binding site"
FT complement (1079..1087)
FT /tag= d
FT /note= "Alfinl binding site"
FT complement (1155..1160)
FT /tag= e
FT /note= "Alfinl binding site"

Db 1342 CAATTGCGATTATGATTTAGATTGCTGCAATACGGTCGCGTAATGTCATCACT 1401
 QY 1403 CACGAGAAAGAGGTATCAAAATTTCAAGGTATTTTATTTATTTTAAACAATAAATTTTC 1462
 Db 1402 CACGAGAAAGAGGTATCAAAATTTCAAGGTATTTTATTTTATTTTAAACAATAAATTTTC 1461
 QY 1463 AAGGCTCTTGTTCACCATATAAACCCTTCACTCACACCAATTCCTTAAAGTGTATGACT 1522
 Db 1462 AAGGCTCTTGTTCACCATATAAACCCTTCACTCACACCAATTCCTTAAAGTGTATGACT 1521
 QY 1523 TCATAGTACACTACTCTTCTTTGAACATGCTTAACATGCTCTAGCCAAATGTTTT 1582
 Db 1522 TCATAGTACACTACTCTTCTTTGAACATGCTTAACATGCTCTAGCCAAATGTTTT 1581
 QY 1583 CATCCTCTCTTGAACATGCTTAACATGCTCTAGCCAAATGTTTT 1613
 Db 1582 CATCCTCTCTTGAACATGCTTAACATGCTCTAGCCAAATGTTTT 1612

RESULT 2
 ID ABZ10246/c
 ID ABZ10246 standard; DNA; 8056 BP.
 XX ABZ10246;
 AC ABZ10246;
 XX
 DT 16-JAN-2003 (first entry)
 XX
 DE Haematopoietic cell proliferation disorder related DNA sequence #386.
 XX
 KW Human; haematopoietic cell proliferation disorder; cytostatic;
 KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KW cytosine methylation state; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277272-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 26-MAR-2002; 2002WO-EP03401.
 XX
 PR 26-MAR-2001; 2001US-278333P.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX

PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T;
 PI Pelet C, Schwöbe I, Ziebarth H;
 XX WPI; 2003-018942/01.
 DR
 XX
 XX Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent
 PT that distinguishes between methylated and non-methylated CpG
 PT dinucleotides -
 XX
 PS Claim 28; SEQ ID 386; 117pp; English.
 XX

CC The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder

CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related
 CC DNA sequences. The nucleotide sequences from the present invention can
 CC also be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables
 CC a highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients.
 XX
 SQ Sequence 8056 BP: 3711 A; 0 C; 371 G; 3974 T; 0 other;
 Query Match 7.1%; Score 114; DB 25; Length 8056;
 Best Local Similarity 44.6%; Pred. No. 5.7e-07;
 Matches 613; Conservative 1; Mismatches 751; Indels 9; Gaps 4;
 QY 1 TTTTATAATATTTAAAGCTTGTATTAATTTTGGGATCTATATATAAGCCCACTACCAAT 60
 Db 1855 TTTTATTTTATTTTAAACCAAAATTTTATTTATTTAAATTAATTAATAATA 1796
 QY 61 TTTAAAT 120
 Db 1795 ACATAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1736
 QY 121 TTTAAAT 180
 Db 1735 TTTAAAT 1679
 QY 181 TTACCTTCTGATGCTAAACCAAAATTAATTTTGTACCAATTTTAAAGTCATATAATAGT 240
 Db 1678 TAAATTTTATCATATTAATTAATAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1619
 QY 241 ACAATCCZACCCCTTTGAGAGGTTAATGTGTGCGGATTTTCTAGATAAACAGGCGCA 300
 Db 1618 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 1559
 QY 301 TTCACGATCTCTTGTGTCACGCTTGGAGAACCCCTATCTCGGCTTGGAGATTTACTTC 360
 Db 1558 TAAAAAATAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1499
 QY 361 TTGTTGATGCTCTAGAGTACAGCTCCTTAAGGCTGTAGTCTAGTCTTTTCTCTCTCT 420
 Db 1498 ATTAATTTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTTATTT 1439
 QY 421 TCCTACCAAAAAAAGTCATATAATAGTCTATATACATATAAATTTTAAATTAATAATAA 480
 Db 1438 TTTTAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 1379
 QY 481 AAAAAATTCATCCCTAAACACATAGTAGAAATTTTCATAAAAAAATAATTTGTTTATAAT 540
 Db 1378 AATAAATTTTAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 1319
 QY 541 ACATCCGCTTACGGTAAACCAAAATGGAATTTGGGTATGGGTATGAGTACTAGTAAATTA 600
 Db 1318 TTTTATTAATTTTAAATTTTAAATAATATATACATTTTATATATATTAATTAATAATA 1259
 QY 601 TCATTGGTTAAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 660
 Db 1258 TAACATTTTATTTTATTCAAAAAATAATTTATTTATTTAATAAAAAAATAAAAAAATA 1199
 QY 661 TTGGAACATGAAGGGTATTTGATTTTACACCTTTTACACCTTTTACACCTTTTACACCT 720
 Db 1198 TTTAAATTTTAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 1139
 QY 721 TCAATATATATTTTGGGGATCAACACCAACCAATCATTTACCATACATGCTTATACATA 780
 Db 1138 TCAAAAAAATAATTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 1079
 QY 781 CCCCGCTCAATCTCTTTTACCAATTAACATTTGAAATGCTTCTCTCTCTCTCTCTCT 840
 Db 1078 TAAACAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1019
 QY 841 ATAAAAACATCAAGCTAGCAAAATGTTGTTTTTGGCATGACATTTTCATATAGTTTA 900

1018	Db	TTTTTAAATAAATATATTTTTTAAAAAAATATAAT- - - ATATATTTTATTCATTTAAATA	963
901	Qy	AAGGATCGATGATTCGATTACAAAAACAAAATACTATAATAATCTAGCACAAAGTTTAAAG	960
962	Db	AAAAAATAAAATTTATTTTTTAAAAAAATTTATACAAAAACAACAAAATAATTAATA	903
961	Qy	CAATATTTAAAGCTTCATAGCATGTGGATATTCATTTAGAAATATAGATTAGATGCC	1020
902	Db	TATAAATTTTAAATAATTTACAAAATTTTTTATACATAAAAAAAATAAAAAACAATTTTT	843
1021	Qy	CTTTTCATCCGGGTCTAACAGACCACCTTCCTACTACATGCTCAAAAATGCTCTAGTAC	1080
842	Db	TTTTAATAAAATAATTTTAAATATCATTTATAAATTTATTTTAAATATATAATTAAT	783
1081	Qy	AGCACCGCTTTTTTACTTGAATCCCTTGTCATGCATGAAAAAAATCAAAACAATATTG	1140
782	Db	AATTTTATAATTTTAAATTTAAATTTTAAAAATTTAAAAATTTAAATATAAATATATT	723
1141	Qy	GACACAAACTTGCCTCCACTTTCCTTTTCTCTGCCCTAGTTGTGTTGAGACTCAT	1200
722	Db	TAAATTTAAATTTTAAATATAAAAAAATTTTTTATATTTATTTTTTTTTTTTTTTTT	663
1201	Qy	ATTGA-TCAAATTTGGCTATGAATTCAAACAAAAATTCCTACTACCCATTCGATGTG	1259
662	Db	ATAAATTTATATTTAAAAATTTATTTTAAAAAATTTTAAAAATTTATATAATTTTTTT	603
1260	Qy	GGGCCACATATAAATCCATGAAGGATTTTCAAT-GTCCATCCCAAGTCAATGATTCACAT	1318
602	Db	TAAATATTATTTTAAATATAAATATTTTAAATAAACATTTTAAATAATTTATTAAT	543
1319	Qy	ATATACATTTGAATAATTTTAAATCCAAATTTGCAGTATTATGATTTAGATTTGATT	1372
542	Db	TTTTAAATTTTAAATAATAATTTTAAATATTTTAAATATATAAATATATAAATTTTAA	489

RESULT 3

ABL70376/C

ID ABL70376 standard; DNA; 16258 BP.

AC ABL70376:

DT 01-JUL-2002 (first entry)

Chemically treated cell signalling pDNA sequence complementary to #133.

Cell signalling: cytosine methylation: cell signalling disease:

KW cancer; tumour; cytostatic; ds.

OS Unidentified.

PN WO200202807-A2.

PD 10-JAN-2002.

29-JUN-2001: 2001WO-EP07471.

30-JUN-2000: 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-154758/20.

Nucleic acid, useful for diagnosis and therapy of diseases associated with cell signalling e.g. cancer, comprises chemically modified genomic sequences of genes associated with cell signalling -

PS	Claim 1:	SEQ ID NO 266:	24 pp+sequence listing:	English:
PS	Claim 1:	SEQ ID NO 266:	24 pp+sequence listing:	English:

CC The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated

with cell signalling. The activity of the modified sequences of the invention may be described as cytostatic. The object of the invention is to provide the chemically modified DNA of genes associated with cell signalling, as well as oligonucleotides and/or PNA-oligonmers for detecting cytosine methylations, as well as a method which is particularly suitable for the diagnosis and/or therapy of genetic and epigenetic parameters of genes associated with cell signalling. The chemically modified DNA provided by the invention is useful for diagnosis and therapy of diseases such as solid tumours and cancer. The sequences given in records ABL70111-ABL70636 represent chemically pre-treated genomic DNA's of genes associated with cell signalling.

Note: The sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.

Sequence 16258 BP; 4242 A; 330 C; 3801 G; 7885 T; 0 other;

Query Match	6.2%	Score 100.6;	DB 24;	Length 16258;
Best Local Similarity	49.0%	Pred. No. 2.9e-05;		
Matches 381; Conservative		0; Mismatches 389;	Indels 7; Gaps 4;	

QY	22	ATAATAATTTGGGATCTTATATATAGACCCACCTACCAATTTAAATTTATATATATATATATA 81
Db	2382	ATAATATATATAATATATATTTCTACCTATATATAAAAAATATATATATATATA 2323
QY	82	TATATATATATATATATAATAATTTTATTTACCAATTTAAATTTATATATATATATA 141
Db	2322	TA-ATATATATATAAATATATATATAATATATATATATATATATATATATA 2264
QY	142	TATATATATATATATATAATAATTTTATTTATTTATTTAGCTTGATGGTAAAAAAT 201
Db	2263	TATACATATATATAAAAAATAAAATATAAAATATATATATAAAAAATATATAAAAAATAT 2204
QY	202	AAATATAATTTGTTACCATTTAAAGTCATAAATATAGTACA-ATCCAAACCCCTTTGGAG 260
Db	2203	ATATATATACTAAAAACATATAAAAAATATAAAATATATAATATATATACATATAAAAAAC 2144
QY	261	GTTAAATGTGTGCGGATTTCTAGATAAACAAAGGYGCCATTCACGANTTCCTTGGTGC 320
Db	2143	ATATATACGTATATATATATATATATATACATTTTAAAAACGTATATATATATATAT 2084
QY	321	AGCTTGGAGAACCTATCCTGGCTTTGGAAGATTTTACTTCTTGTGATGCTTCTAGA--- 377
Db	2083	ATAATATATACATATAAAAAAACGTATATATATATAATATTTATATACATATAAAAA 2024
QY	378	-GTACAGCTCCTTAAAGCGTAGCTAGTTTTTTTTTTTTCATCCTCTCCTACCAAAAAAAA 436
Db	2023	TATATATATAATATATATACCATAAAAAATATATATATTTATTTATACATAAAAAACGTA 1964
QY	437	AAAGTCATAAATAGTTTTATACATATAACTTTTAATAAAAAATAAAAAATTTTCATCCCTA 496
Db	1963	TATATAATATATATAATATATATATATATATATATATATAATATATAATATATATATA 1904
QY	497	AAAAACATAGTAGAAATTTTCATAAAAAAATAT-TGTTTATAAATTTACATGCCGTTACGGT 555
Db	1903	TATAATATATATTTATATATATATAAATATATATATATATATATATATATATATATAA 1844
QY	556	AAAAAATGATAAATTTGGGTATCGGACTAGTAATTAATAAGGTCATTTGGTTAAAAAAA 615
Db	1843	TATATATATAATATATATATTTATATATATATAAATATATACTATATATATATATAAAT 1784
QY	616	ACTAAAAAATAATTTCTCCTCATTTATATGAATGACATTTTTTTTGGAAACATGAAGGG 675
Db	1783	ATATATATAATATATATATATATATATATATATATATATATATAATATATATTTTTACTCTCA 1724
QY	676	TATTTGATTTTTTACCACCTTTTACACCTTTTCAAAAGCCATTCGAAGGATGAATATAGATTTTT 735
Db	1723	CCAAAACCTCTTACAAATACTACTCTTCAAAAATTTCCATCCACCCCATCTAAAAACCCCAT 1664
QY	736	GGCGGATCAACACAGAATCATACGATAACATGCTTTATACATACCCCGTCAATCT 792
Db	1663	AACAAAACAAAAATAATCAATATTTCTCTCCCTTTTCTAACTCCCTCACTTACCTTCT 1607

XX (EPIC-) EPIGENOMICS AG.
 PA Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T;
 PI Pelet C, Schwöpe I, Ziebarth H;
 XX WPI; 2003-018942/01.
 DR
 XX
 XX Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent
 PT that distinguishes between methylated and non-methylated CpG
 PT dinucleotides -
 XX
 XX Claim 28; SEQ ID 386; 117pp; English.
 XX
 CC The present invention describes a method for detecting and
 CC differentiating between hematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy hematopoietic cells and proliferative
 CC disorder hematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of hematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of hematopoietic cell proliferation disorder related
 CC DNA sequences. The nucleotide sequences from the present invention can
 CC also be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC hematopoietic cell proliferative disorders. The present method enables
 CC a highly specific classification of hematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients.
 XX
 SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 other;

XX Query Match 6.2%; Score 100.2; DB 25; Length 8056;
 XX Best Local Similarity 46.8%; Pred. No. 3.6e-05;
 XX Matches 472; Conservative 1; Mismatches 519; Indels 21; Gaps 4;
 QY 1 TTTTATAAATTTTAAAGCTTGATATAAATTTTGGCGATCTATATAAGCCCACTACCAAT 60
 DB 1010 TTTTAAAAATTTTATATATTTTAAATTTTGAATTTAAATATGCTAAATAAATAAT 1069
 QY 61 TTAATAATTAT 120
 DB 1070 TTTGTTTAAATTTGTTTTTTTTTAAATGTTTTAAATGTTATATATTTCTTTTAAATAT 1129
 QY 121 TTAATAATTATTA 180
 DB 1130 TTTTGTGATTTGAAAAATTTAAATTAATATATATTTTGTATTTATTAATTAATTTTAA 1189
 QY 181 TTACGTTGATGTTAAAAATTAATATATATTTTGTACCAATTTTAAAGTCATATAATAGT 240
 DB 1190 AAATTTTATATATTTTAAATTTTAAATTAATAAATAAATTTTTCGATAAATAA 1249
 QY 241 ACAATCCAAACCCCTTTGAGAGGTTAATGTGTGCGGATTTTCTAGATAAACAAGYGCCA 300
 DB 1250 AAAATGTTATGTAATATATATATATATATATATATATATATTTTAAAAATTAATAAT 1309
 QY 301 TTCACGATTTCTTCTGGGCGACCTTGGAGAACCCCTATCGCTGGCTGGAGATTTACTTC 360
 DB 1310 TATAAAAAAATTTATATATATTTTAAATAAATAAATAAATTTTAAATTTTATTTTAA 1369
 QY 361 TTGTTGATCTCTAGAGTACAGCTCCCTTAAAGGCTGATGCTAGTCTAGTCTTTTTCATCCT 420
 DB 1370 AAATTTATTTTTTTTTTATATAAATAATATGATTTTTTTTTTATTTTTTTTTTTTTT 1429

QY 421 TCCTACCAAAAAA-AAAAGTCATAAATATAGTTTTATACATATAAATTTTAAATAAATA 479
 DB 1430 TTTTAAAAAATAAATAAATTTTATTTTAAATAAATATATATTAATAATTAATATAT 1489
 QY 480 AAAAATTTTCATCCCTTAAAAACATAGTAGAAATTTTCATATAAATAAATATGTTTATAAT 539
 DB 1490 TAAATTAATATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 1549
 QY 540 TACATGCCCTTACGGTAAATAATGGATAAATTTGGGTATGGAGTACTAGTAAATAAGG 599
 DB 1550 -----ATTTTTTAAAAAATATATTTAAATTTAAATTTTAAATTTTATTTATAAT 1602
 QY 600 TTCAITGGTTAAAAAACTAAAAATAATTTCTCTCTGATTTATATATATAGAAATGACATTTT 659
 DB 1603 TTAATATTTTAAAAA- -TTGAATAAATGAATGTAATAAATAAATAAATAAATAAAT 1659
 QY 660 TTTGGAACATGAAGGTATTTGATTTTACCACCTTTTACACCTTTTCAAGCCATTTCAAGG 719
 DB 1660 TTAATATGATAAATTTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAT 1719
 QY 720 ATGAATATAGATTTTGGCGCATCAACACACAGAATCATTCAGATAACATGCTTATACAT 779
 DB 1720 ATATTTATATTTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAAT 1779
 QY 780 ACCCGCTCAATCTCTTTTTCACCAATAAACAATTTGAAATGTTGCTCTTTTCGTTAAG 839
 DB 1780 TAAAAATTAATATGTTATATATTTTAAATAAATAAATAAATAAATAAATAAATAAATA 1839
 QY 840 CATAAAAACATCAAGCTCTAGCAAAATGTTGTTTTGCGATGACACATTTTCATATAGTT 899
 DB 1840 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1889
 QY 900 AAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959
 DB 1890 TTAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1949
 QY 960 GCAATATTAAGAAGTTTCATAGCATGTTGATATTCATTTAGAAATATAGATTA 1012
 DB 1950 AAAATATATATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2002

RESULT 6
 ABZ10100/c
 ID ABZ10100 standard; DNA; 8056 BP.
 XX
 AC ABZ10100;
 DT 16-JAN-2003 (first entry)
 XX
 DE Haematopoietic cell proliferation disorder related DNA sequence #240.
 XX
 KW Human; haematopoietic cell proliferation disorder; cytostatic;
 KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KW cytosine methylation state; gene; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200277272-A2.
 XX
 PD 03-OCT-2002.
 XX
 XX 26-MAR-2002; 2002WO-EP03401.
 XX
 XX 26-MAR-2001; 2001US-278333P.
 XX
 XX (EPIC-) EPIGENOMICS AG.
 XX
 XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T;
 PI Pelet C, Schwöpe I, Ziebarth H;
 XX

DR WPI; 2003-018942/01.
 XX Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent
 PT that distinguishes between methylated and non-methylated CpG
 PT dinucleotides -
 XX
 PS Claim 28; SEQ ID 240; 117pp; English.
 XX
 CC The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. AB209861 to AB211118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related
 CC DNA sequences. The nucleotide sequences from the present invention can
 CC also be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables
 CC a highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients.
 XX
 SQ Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 other;
 Query Match 5.9%; Score 95.2; DB 25; Length 8056;
 Best Local Similarity 44.4%; Pred. No. 0.00016;
 Matches 551; Conservative 1; Mismatches 679; Indels 9; Gaps 4;
 1 TTTTATAATATTAAGCTTCATATATTTGGCATATATATAAGCCACACCAAT 60
 DB 1855 TTTTATTTTATTTTAAACGAAATTTATTTATTTATTAATTAATATATA 1796
 61 TTAAT 120
 DB 1795 ACGTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTA 1736
 121 TTAAT 180
 DB 1735 TTAAT 1679
 181 TTACGTTGATGTTAAATATATATATATATATATATATATATATATATAT 240
 DB 1678 TAAATTTTATCGTATTAATTAATTAATTTTATTTTATTTTATTTTATTT 1619
 241 ACAATCCACCCCTTGAGAGGTTAATGTTGGCGGATTTCTAGATAAACAAGYGCCA 300
 DB 1618 TTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1559
 301 TTCAGGATCTCTTGGTGGAGCTTGGAGACCCCTATCCTGGCTTGGAGATTTACTTC 360
 DB 1558 TAAATATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1499
 361 TTGTTGATGCTCTAGAGTACAGCTCCCTTAAGGCTGTAGTCTAGTTTTTTTTCATCCT 420
 DB 1498 ATTAATTTAATTAATTTAATTAATTTAATTTAATTTAATTTAATTTT 1439
 421 TCCTACCAAAAAAAGATCAATAATATAGTTTATACATATACATTTAATAAATAA 480
 DB 1438 TTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1379
 481 AA-AAATTTTCATCCCTTAAACATAGTAGAATTTTCAATAAATAAATAATTTGTTATAAT 539
 DB 1378 AATAAATTTTAAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTAAT 1319

QY 540 TACATGCGTTACGGTAAAAAATGGATAAATGGGTATGGTACTAGTAAATTAATAAGG 599
 DB 1318 TTTTATAATTTATTAATTTTAAATAATATACGTTTATATATATTAATAATTAACG 1259
 QY 600 TTCATTGGTTAAAAAATAAATAATTTCTCTCGATTTATATGAATGACATTTT 659
 DB 1258 TAACGTTTATTTTATTCGAAAAATTAATTTATTAATTAATAAATAAATAAATAA 1200
 QY 660 TTTGGAACATGAAGGTATTGATTTTACCACCTTTTACACCTTTCAAGCCATTCACAGG 719
 DB 1199 ATTAATAATTTTAAATAAATAATTAATAATCGAAATATATTAATTTTAAATTTTCGGAAT 1140
 QY 720 ATGAATAATAGATTTTGGCGCATCAACACAGAATCATTCAGATAAATGCTTATACAT 779
 DB 1139 ATCGAAAAAATAATTAATAACGATATATAACGTTTAAACGTTTAAAAAATAACGNA 1080
 QY 780 ACCCGGTCAATCTCTCTTTTACCAATTAACATTAAGATTTGCTCTCTTCGTTAAG 839
 DB 1079 TTAACACGAAATTTTATTTTACGTAATTAATTTTAAATTTTAAATAAATAAATAA 1020
 QY 840 CATAAAAATCAAAAGTCTAGCAAAATGTTGTTGGCATCACACATTTTCATATAGTTT 899
 DB 1019 ATTTTAAATAATTAATTTTAAATAAATAATTAATTTTATTTTATTCGTTTAAAT 964
 QY 900 AAAGGATCATGATTCGATTTACAAAAAATACTAATAATCTAGCACAAAGTTTAAA 959
 DB 963 AAAAAAATAAATAATTTTAAATAAATAATTTATTCGAAAAACGCAATAATTAATA 904
 QY 960 GCAATATATTAAGCTTCATAGCATGTGATATTCATTTAGAAAATATAGATAGATGCG 1019
 DB 903 ATATAAATTTAAATAATTTACGAAATTTTATACGTAAAAAATAAATAAATAAATAA 844
 QY 1020 CCTTTCATCAGGGTCTAACAGCACCCTGTCACATACATGTCAAAAATGTCCTCTAGTA 1079
 DB 843 TTTTAATAAATAATTTTAAATAATTCGTTTAAATAATTTTAAATAATTAATAATAATA 784
 QY 1080 CAGCACCCGCTTTACTTGTGATTCCTGTCATGCATGCAAAAAATAAATAAATAAATA 1139
 DB 783 TAAATTTTATAATTTTAAATAATTTTAAATAATTTTAAATAATTAATAATAATAATA 724
 QY 1140 GGACACACAACTTCCGCCCTGCTTTCTTTCTTCTGCCCCCTAGTTTGTGTTGACATCA 1199
 DB 723 TTAATTAATAATTTTAAATAAATAAATAATTTTATATATTTATTTATTTTATTTT 664
 QY 1200 TATTCATCAATTTGGCTATGAATTCACAAACAAAAATTCATCA 1239
 DB 663 AATAAATTTAATTTTAAATAATTTTAAATAATTTTAAATAATTTTAA 624
 RESULT 7
 ABL54336
 ID ABL54336 standard; DNA; 6641 BP.
 XX
 AC ABL54336;
 XX
 DT 29-JUL-2002 (first entry)
 XX
 DE Chemically treated apoptosis gene complementary to gene #18.
 XX
 KW Apoptosis; HIV; Bloom syndrome; cardiopathy;
 KW neurodegenerative disorder; Herpes simplex virus; renal ischaemia;
 KW amyotrophic lateral sclerosis; cancer; ds.
 XX
 OS Unidentified.
 XX
 PN WO200177164-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-EP03969.
 XX
 PR 06-APR-2000; 2000IE-1019058.

PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-017444/02.
 XX
 XX Chemically modified sequences of genes associated with apoptosis are
 PT useful to determine methylation patterns of genomic DNA samples for
 PT diagnosis of associated diseases such as cancer
 XX
 PS Claim 1; Seq ID #36; 24pp; English.
 XX
 CC This invention relates to chemically pre-treated DNA of genes
 CC associated with apoptosis. The nucleic acids are used to allocate
 CC patients for specific therapy for HIV infection, Bloom syndrome,
 CC cardiopathy, aging, neurodegenerative disorders, Herpes simplex virus
 CC infection, renal ischaemia, amyotrophic lateral sclerosis, solid tumours
 CC and cancers. This nucleotide sequence represents a chemically
 CC treated apoptosis gene. Even SEQ ID numbers are the complementary
 CC DNA strands to the odd SEQ ID numbers. The sequence data for this
 CC patent is not represented in the printed specification but is based on
 CC information supplied by the European patent office.
 XX
 SQ Sequence 6641 BP; 2495 A; 22 C; 984 G; 3140 T; 0 other;

Query Match 5.88; Score 93.2; DB 24; Length 6641;
 Best Local Similarity 46.58; Pred. No. 0.00031;
 Matches 451; Conservative 0; Mismatches 503; Indels 16; Gaps 4;

QY 12 TTTAAGCTTCGATTAATTTGGCGATCTATATATA-AGCCCACTACCAATTTAAATAT 70
 DB 2349 TATAATTTAGATTTTATTAGGAGATATATATATATATATATATATATATATAT 2408

QY 71 AT 130
 DB 2409 AT 2468

QY 131 AT 190
 DB 2469 AT 2525

QY 191 GGTAAAAAATAATATAATTTGTTACCAATTTAAAAAGTCATAAATAGTACAAATCAAC 250
 DB 2526 TAT 2585

QY 251 CCTTTGAGAGGTTAATGTGTGCGGATTTCTAGATAAACAGGYGCCATTCACGATTC 310
 DB 2586 TAT 2645

QY 311 TTTCTTGGTCGAGCTTGGAGACCCCTATCCTGGCTTGGAGATTTACTTCTGTTGATGC 370
 DB 2646 TTAATAAAGAGATATGATTTAAGTAGGTTGTATGTGGAAAGCTTAGATATTTTT-T 2701

QY 371 TTTCTAGAGTACACTCCTTAAGGCTGTAGTCTAGTTTTTTTTTTTTCATCCTTCTACCAA 430
 DB 2702 TTTTTTTGTTAATTTTAAAGATTAATTAATTTAGTTTATTAATTAAGGATAGTTTTT 2761

QY 431 AAAAAAAGTCATAAATAGTTTTATACATATAACTTTTAATAAAAAATAAAAAAATTTCA 490
 DB 2762 ATATAAAATATATTTTTTAAAGAAATA-----TGTAGGATATGAATAAAATTTT 2813

QY 491 TCCCTAAAAACATAGTAGAATTTTCATAAAAAAATATTTGTTTATATTTACATGCCGTT 550
 DB 2814 GGTTTAGAAGTATAGATTTATTTTTTTAGTGTATTTTGTAGTTTATATGTTAATAATGA 2873

QY 551 ACGGTAAAAAATGGATAATTTGGGTATGGAGTACTAGTAATTAATAAGCTTCATGGTTA 610
 DB 2874 TAATTTATTAGAAGATATATGGAAGTTATAAATAGGATAAAGATATATATTTAATA 2933

QY 611 AAAAAAATAAATAATTTCTCTCGATTTATATGAAATGACATTTTTTTGGACATG 670
 DB 2934 AATGTAATAATTTTTTTATGATTTTTTCTGATTAATAAATGAATTTATATAAAGTAATGAT 2993

QY 671 AAGGTATTCATTTTTTACACCTTTTACACCTTTTCAAGCCATTCAGGATGATATAGA 730
 DB 2994 AATTATATATATATATGAGTAATTTATTTATTTATTTATTTATTTATTTATTTATTTA 3053

QY 731 TTTTGGGCGATCAAAACACAAAGATCATTACGATAACATGCTTATACATACCCCGTCAAT 790
 DB 3054 TTATATATGATTTTATTTAGTAATTTATTTATTTATTTATTTATTTATTTATTTATTTA 3113

QY 791 CTCTCTTTTTTACCCTAAATGAAATGTTGCTTTCTGTTAGCATAAAAACAT 850
 DB 3114 ATTATTATATGATATAATAAATAGTAATTTATTTATTTATTTATTTATTTATTTATTTA 3173

QY 851 CAAAGTCTAGCAAAATGTTGTTTTTGGGATGACACATTTTCATATAGTTTTTAAAGATGCAT 910
 DB 3174 TTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 3233

QY 911 GATTGATTCACAAAAACAAATACTAATAATTTCTAGCACAAAGTTTAAAGCAATATTTATA 970
 DB 3234 TTAAGTATTTTATAAATAGTAGTATATTTGATTTTATTTATTTATTTATTTATTTATTTA 3293

QY 971 AAGCTTCATA 980
 DB 3294 TATTTTATTA 3303

RESULT 8
 ABL32315
 ID ABL32315 standard; DNA; 6641 BP.
 XX
 AC ABL32315;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 288.
 DE
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; anti-naemic; cytosine; neurotropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antichronic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-130909/17.
 DR
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 XX
 PS Claim 1; SEQ ID NO 288; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences

CC amplicates carry a detectable label. The method further involves
 CC identifying methylation status of one or more cytosine positions, and
 CC analysing methylation status of the cytosine positions by reference to
 CC one or more data sets. The genomic DNA is chemically treated by using a
 CC bisulphite, hydrogen sulphite or disulphite. The amplification
 CC step amplifies DNA which is of particular interest in astrocytoma or
 CC brain tissue, based on the specific genomic methylation status of brain
 CC tissues, as opposed to background DNA. The amplicates carry a
 CC fluorescent label or radionuclide. Optionally, the labels of the
 CC amplicates are detachable molecule fragments having a typical mass
 CC which are detected in a mass spectrometer. The fragments of chemically
 CC pre-treated genomic DNA to be amplified, have a single positive or
 CC negative charge for a better detectability in the mass spectrometer.
 CC Preferably, the amplicates or fragments of the amplicates are
 CC detected by matrix assisted laser desorption/ionization mass spectrometry
 CC (MALDI) or using electron spray mass spectrometry (ESI). The
 CC present sequence is one of the chemically pre-treated reference DNA
 CC samples of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 6621 BP; 1770 A; 214 C; 1781 G; 2856 T; 0 other;

Query Match 5.7%; Score 92; DB 24; Length 6621;
 Best Local Similarity 61.9%; Pred. No. 0.00044;
 Matches 146; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 2 TTTTAAATTTTAAAGCTTGATATAATTTTGGCATCTATATATAAGCCCACTACCAATT 61
 Db 4889 TGTGTGATATATAATAATATATATATATGTTGATATATATAATATATATATATG 4948
 QY 62 TAAATTTAT 121
 Db 4949 TGTA 5008
 QY 122 TAAATTTAT 181
 Db 5009 TAT 5068
 QY 182 TACGTTGATGTAAGTAAATAATATATATATATATATATATATATATATATATATAT 237
 Db 5069 GTAAT 5124

RESULT 12

AAS61098 standard; DNA; 6621 BP.

XX AAS61098;

XX 29-JAN-2002 (first entry)

DE Human gene regulation-associated gene oligonucleotide #53.

XX Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preclampsia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
 KW immunostimulant; cardiant; antiinflammatory; coagulant; antilasthmatic;
 KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.

OS Homo sapiens.

XX WO20017375-A2.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-EP03968.

XX 06-APR-2000; 2000DE-1019058.

PR

07-APR-2000; 2000DE-1019173.

PR

30-JUN-2000; 2000DE-1032529.

PR

01-SEP-2000; 2000DE-1043826.

XX

(EPIG-) EPIGENOMICS AG.

PA

Olek A, Piepenbrock C, Berlin K;

XX

WPI; 2002-017470/02.

XX

PT New nucleic acid sequences from chemically modified genes associated
 PT with gene regulation, useful for analysing cytosine methylations for
 PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
 PT disease

XX

PS Claim 1; SEQ ID No 54; 26pp; English.

XX

CC The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The
 CC chemical pretreatment converts cytosine bases unmethylated at the
 CC 5-position to uracil or another base with hybridisation behaviour
 CC dissimilar to cytosine, to enable analysis of cytosine methylations.
 CC The DNA sequences, oligomers (or sets/arrays) and method are
 CC useful in the diagnosis of diseases (or predisposition to diseases)
 CC associated with gene regulation and in therapy of such diseases, by
 CC enabling analysis of the cytosine methylation patterns of such genes,
 CC kits are provided. They are especially useful in diagnosis
 CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
 CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
 CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
 CC preclampsia, graft versus-host disease. The present sequence is a
 CC sequence included in the sequence data for this specification and is
 CC associated with the human gene regulation-associated genes.

CC

CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX

XX Sequence 6621 BP; 1770 A; 214 C; 1781 G; 2856 T; 0 other;

XX

XX

QY

Query Match 5.7%; Score 92; DB 24; Length 6621;
 Best Local Similarity 61.9%; Pred. No. 0.00044;
 Matches 146; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY

2 TTTTAAATTTTAAAGCTTGATATAATTTTGGCATCTATATATAAGCCCACTACCAATT 61

Db

4889 TGTGTGATATATAATAATATATATATATGTTGATATATATAATATATATATATATATATG 4948

QY

62 TAAATTTAT 121

Db

4949 TGTA 5008

QY

122 TAAATTTAT 181

Db

5009 TAT 5068

QY

182 TACGTTGATGTAAGTAAATAAT 237

Db

5069 GTAAT 5124

XX

16-JAN-2003 (first entry)

XX

Haematopoietic cell proliferation disorder related DNA sequence #124.

XX

Human; haematopoietic cell proliferation disorder; cytostatic;

XX

ABZ09984 standard; DNA; 6621 BP.

XX

ABZ09984;

XX

16-JAN-2003 (first entry)

XX

Haematopoietic cell proliferation disorder related DNA sequence #124.

XX

Human; haematopoietic cell proliferation disorder; cytostatic;

THIS PAGE BLANK (USPTO)


```
; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801.861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-3

Query Match
Best Local Similarity 4.7%; Score 75.6; DB 4; Length 53332;
Matches 138; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 2 TTTATAAATATTTAAGCTTGATATAATTTTGGCATCTATATATAAGCCCACTACCAATT 61
DB 31616 TATATAAATATATAAATATATAAATATAAATATAAATATAAATATAAATATAA 31557

QY 62 TAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 121
DB 31556 TATATAAATATATAAATATAAATATAAATATAAATATAAATATAAATATAATAT 31497

QY 122 TAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 181
DB 31496 AATATATATATAAATATATATATAAATATAAATATAAATATAAATATAAATAT 31437

QY 182 TACGTTGATGCTAAATAAATATAATTTTGGTACCAATTTTAAAGTCATAAATATAGTA 241
DB 31436 ATAAATATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAA 31377

QY 242 CA 243
DB 31376 TA 31375

RESULT 3
US-09-783-203-1/c
; Sequence 1, Application US/09783203
; Patent No. 6576464
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Lebkowski, Jane
; TITLE OF INVENTION: Tpacked stem cells
; FILE REFERENCE: 096/003
; CURRENT APPLICATION NUMBER: US/09/783.203
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/253,443
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 15418
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-783-203-1

Query Match
Best Local Similarity 4.7%; Score 75.4; DB 4; Length 15418;
Matches 136; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 2 TTTATAAATATTTAAGCTTGATATAATTTTGGCATCTATATATAAGCCCACTACCAATT 61
DB 1185 TTTATATATGATATATTTGATATATCTGATATATATATATATATATATATATATA 1126

QY 62 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 121
DB 1125 TTTATATATGATATATATATATATATATATATATATATATATATATATATATATAT 1066

QY 122 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 181
```

```
DB 1065 TATATATGATATATATATATATATATATATATATATATATATATATATATATATAT 1006
QY 182 TACGTTGATGCTAAATAAATATAAATTTTGGTACCAATTTTAAAGTCATAAATATA 238
DB 1005 TATGTAATATATATATATATATATATATATATATATATATATATATATATATATAT 949

RESULT 4
US-09-783-203-1
; Sequence 1, Application US/09783203
; Patent No. 6576464
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Lebkowski, Jane
; TITLE OF INVENTION: Tpacked stem cells
; FILE REFERENCE: 096/003
; CURRENT APPLICATION NUMBER: US/09/783.203
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/253,443
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,357
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 15418
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-783-203-1

Query Match
Best Local Similarity 4.6%; Score 73.8; DB 4; Length 15418;
Matches 157; Conservative 0; Mismatches 102; Indels 2; Gaps 2;

QY 2 TTTATAAATATTTAAGCTTGATATAATTTTGGCATCTATATATAAGCCCACTACCAATT 61
DB 979 TATATACATATATAAATATATATTTACATATATAAATATATACATATATAAATA-TACATATA 1037

QY 62 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 121
DB 1038 TAAATACATATATAAATATATACATATATAAATATATACATATATAAATATACATATA 1037

QY 122 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 180
DB 1098 TATACATATATAAATATATACATATATAAATATATATATATATATATATATATAAATA 1157

QY 181 TTACGTTGATGCTAAATAAATATAAATTTTGTACCAATTTTAAAGTCATAAATATAGT 240
DB 1158 AGTATATACAAATATATACATATATAAATGATATAGTATATACATATATATATAAATA 1217

QY 241 ACAATCCACCCCTTTGAGAGG 261
DB 1218 TATAAAAAAACTTTTGGCTGG 1238

RESULT 5
US-09-671-317-141
; Sequence 141, Application US/09671317
; Patent No. 6528260
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLIS
; FILE REFERENCE: 62.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/671.317
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 09/536.178
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/IB00/00403
```


[illegible]


```

; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 199122..201122
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 201123..201234
; OTHER INFORMATION: exon S
; NAME/KEY: exon
; LOCATION: 201123..201560
; OTHER INFORMATION: exon S2
; NAME/KEY: exon
; LOCATION: 214676..214793
; OTHER INFORMATION: exon T
; NAME/KEY: exon
; LOCATION: 215702..215746
; OTHER INFORMATION: exon U
; NAME/KEY: exon
; LOCATION: 216836..216994
; OTHER INFORMATION: exon V
; NAME/KEY: exon
; LOCATION: 216836..217077
; OTHER INFORMATION: exon V2
; NAME/KEY: exon
; LOCATION: 217671..217764
; OTHER INFORMATION: exon V1
; NAME/KEY: exon
; LOCATION: 227655..227736
; OTHER INFORMATION: exon V4
; NAME/KEY: exon
; LOCATION: 238715..238919
; OTHER INFORMATION: exon V3
; NAME/KEY: exon
; LOCATION: 240440..240673
; OTHER INFORMATION: exon W
; NAME/KEY: exon
; LOCATION: 240440..241153
; OTHER INFORMATION: exon W2
; NAME/KEY: exon
; LOCATION: 241072..241291
; OTHER INFORMATION: exon X
; NAME/KEY: exon
; LOCATION: 244353..244561
; OTHER INFORMATION: exon Y
; NAME/KEY: exon
; LOCATION: 246273..247802
; OTHER INFORMATION: exon Z
; NAME/KEY: misc_feature
; LOCATION: 247803..249803
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 8316
; OTHER INFORMATION: 99-27943-150 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 21672
; OTHER INFORMATION: 99-27935-193 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 65485
; OTHER INFORMATION: 8-128-33 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 95396
; OTHER INFORMATION: 99-31960-363 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 107281
; OTHER INFORMATION: 99-24656-260 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 160640
; OTHER INFORMATION: 99-24639-163 : polymorphic base A or C
; NAME/KEY: allele
; LOCATION: 160876
; OTHER INFORMATION: 99-24634-108 : polymorphic base A or T
; NAME/KEY: allele
; LOCATION: 168974
; OTHER INFORMATION: 99-7652-162 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 170810
; OTHER INFORMATION: 99-16100-147 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 173358
; OTHER INFORMATION: 99-5862-167 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 189957
; OTHER INFORMATION: 99-5919-215 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 197163
; OTHER INFORMATION: 99-24658-410 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 200778
; OTHER INFORMATION: 8-303-235 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 202651
; OTHER INFORMATION: 8-300-221 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 202679
; OTHER INFORMATION: 8-300-193 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 203378
; OTHER INFORMATION: 8-299-128 : polymorphic base A or T
; NAME/KEY: allele
; LOCATION: 204138
; OTHER INFORMATION: 8-296-213 : polymorphic base A or T
; NAME/KEY: allele
; LOCATION: 204605
; OTHER INFORMATION: 8-252-190 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 204934
; OTHER INFORMATION: 99-24644-194 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 205206
; OTHER INFORMATION: 8-295-248 : polymorphic base A or C
; NAME/KEY: allele
; LOCATION: 205329
; OTHER INFORMATION: 8-295-125 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 206064
; OTHER INFORMATION: 8-293-130 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 206545
; OTHER INFORMATION: 8-292-198 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 207313
; OTHER INFORMATION: 8-251-322 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 208285
; OTHER INFORMATION: 8-289-322 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 208960
; OTHER INFORMATION: 8-287-249 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 209123
; OTHER INFORMATION: 8-287-86 : polymorphic base A or T
; NAME/KEY: allele
; LOCATION: 209631
; OTHER INFORMATION: 8-285-319 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 210361
; OTHER INFORMATION: 8-283-278 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 210463
; OTHER INFORMATION: 8-283-176 : polymorphic base A or G
; NAME/KEY: allele

```


Result No.	Query #	Score	Match	Length	DB	ID	Description
1	5.8	93.2	5.8	6641	12	US-10-311-455-288	Sequence 288, Appli
2	5.7	92.4	5.7	9770	12	US-10-311-455-5	Sequence 5, Appli
c	5.7	92	5.7	6621	14	US-10-172-086-10	Sequence 10, Appli
c	5.7	91.6	5.7	11745	12	US-10-240-453-206	Sequence 206, Appli
c	5.7	91.4	5.7	6641	12	US-10-311-455-288	Sequence 288, Appli
c	5.6	91	5.6	3991	14	US-10-074-045-60	Sequence 60, Appli
c	5.6	90.6	5.6	8136	12	US-10-311-455-528	Sequence 528, Appli
7	5.6	90	5.6	11745	12	US-10-240-453-206	Sequence 206, Appli
c	5.5	89.4	5.5	6621	14	US-10-172-086-10	Sequence 10, Appli
10	5.5	89.4	5.5	513509	11	US-09-754-853A-4	Sequence 4, Appli
11	5.5	89.2	5.5	8136	12	US-10-311-455-528	Sequence 528, Appli
12	5.4	87.4	5.4	6065	12	US-10-311-455-478	Sequence 478, Appli
13	5.4	87.4	5.4	14551	12	US-10-240-485-138	Sequence 138, Appli
c	5.4	87.2	5.4	8136	12	US-10-311-455-527	Sequence 527, Appli
c	5.4	87.2	5.4	14551	12	US-10-240-485-138	Sequence 138, Appli
c	5.4	87.2	5.4	18855	12	US-10-311-455-584	Sequence 584, Appli
16	5.4	87.2	5.4	18855	12	US-10-311-455-584	Sequence 584, Appli


```

; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Method and nucleic acids for the differentiation
; FILE REFERENCE:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/172,086
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 10
; LENGTH: 6621
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-172-086-10

Query Match          5.5%; Score 89.4; DB 14; Length 6621;
Best Local Similarity 68.7%; Pred. No. 0.00092;
Matches 123; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 2 TTTATAAATATTTAAAGCTTGATAATAATTTTGGCGATCTATATAAGCCCACTACCAATT 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5100 TATATATATATATATATTTTATATATATATATATATATATATATATATATATATAT 5041

QY 62 TAAAATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 121
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5040 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4981

QY 122 TAAAATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4980 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4922

RESULT 10
US-09-754-853A-4
; Sequence 4, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Haug, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Soybean Cyst Nematode Resistance
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 4
; LENGTH: 513509
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (111805)..(113968), (114684)..(115204)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(513509)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 318013_region_A3
US-09-754-853A-4

Query Match          5.5%; Score 89.4; DB 11; Length 513509;
Best Local Similarity 47.5%; Pred. No. 0.0026;
Matches 296; Conservative 0; Mismatches 326; Indels 1; Gaps 1;

QY 40 ATATATAGCCCACTACCAATTTAAATTTATATATATATATATATATATATATATATATATA 99
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 464549 ATAATAAGCACTAAACATATGAAAAATTTCTCACATCAATTAATGTTTAAATATAC 464508

QY 100 TAATAATTTTATTTACCAATTTAAATTTATATATATATATATATATATATATATATATATA 159
```

```

; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Method and nucleic acids for the differentiation
; FILE REFERENCE:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/172,086
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 10
; LENGTH: 6621
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-172-086-10

Query Match          5.5%; Score 89.4; DB 14; Length 6621;
Best Local Similarity 68.7%; Pred. No. 0.00092;
Matches 123; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 2 TTTATAAATATTTAAAGCTTGATAATAATTTTGGCGATCTATATAAGCCCACTACCAATT 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5100 TATATATATATATATATTTTATATATATATATATATATATATATATATATATATAT 5041

QY 62 TAAAATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 121
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5040 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4981

QY 122 TAAAATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4980 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4922

RESULT 11
US-10-311-455-528
; Sequence 528, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Det
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 528
; LENGTH: 8136
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-528

Query Match          5.5%; Score 89.2; DB 12; Length 8136;
Best Local Similarity 68.1%; Pred. No. 0.001;
Matches 124; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 TTTTATAAATATTTAAAGCTTGATAATAATTTTGGCATCTATATATATATATATATATATAT 60
```


RESULT 14

```
US-10-311-455-527/c
; Sequence 527, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Sequence of a Polynucleotide
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 527
; LENGTH: 8136
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-527
```

```
Query Match          5.4%; Score 87.2; DB 12; Length 8136;
Best Local Similarity 70.7%; Pred. No. 0.002;
Matches 116; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 23 TAATAATTTTCGATCATATATAAGCCACACCAATTTAAATTTATATATATATATATAT 82
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1014 TCATAATTTTACATTTTATATTTAAATCTATATATATATATATATATATATACACAC 955

QY 83 ATATATATATATATATAATAATTTTATTTTACCAATTTAAATTTATATATATATATAT 142
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 954 ACGTATATATATATATATATATATATATATATATATACACATATATATATATACACATAT 895

QY 143 ATATATATATATATATAATAATTTTATTTATTTATTTATTTTACGTT 186
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 894 ATATATATATATATATATATATTTTATTTTATTTTATTTTCT 851
```

RESULT 15

```
US-10-240-485-138/c
; Sequence 138, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240.485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 138
; LENGTH: 14551
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

; FEATURE:

```
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-138

Query Match          5.4%; Score 87.2; DB 12; Length 14551;
Best Local Similarity 66.5%; Pred. No. 0.0023;
Matches 125; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 58 AATTTAAATTTATATATATATATATATATATATATATATATATATATATATATATAT 117
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2910 AATTCAAATATATATATATATATATATATATATATATATATATATATATATATATA 2851

QY 118 AATTTAAATTTATATATATATATATATATATATATATATATATATATATATATATAT 177
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2850 TATATATATATATATATATATATATATATATATATATATATATATATATATACAAAC 2791

QY 178 TTATTACGTTTCATGCTAAATAAATAATTAATTTGTTACCAATTTAAAGTCATAAATAT 237
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2790 CTAATAACCTAAATAAATAAATCTTTAATCAAAATATATATATTTTCAACATTCATCATTAA 2731

QY 238 AGTACAAT 245
   | | | |
DB 2730 ACAAATAT 2723
```

Search completed: August 14, 2003, 23:27:34
Job time : 439 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 16:08:55 ; Search time 3661 Seconds

(without alignments)
10708.309 Million cell updates/sec

Title: US-09-647-841B-1

Perfect score: 1613

Sequence: 1 ttttataatatttaagctt.....tgaacttgagtaccttactc 1613

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_estc:**

9: gb_est1:**

10: gb_est2:**

11: gb_est3:**

12: gb_est4:**

13: gb_est5:**

14: gb_est6:**

15: em_estfun:**

16: em_estom:**

17: em_gss_hum:**

18: em_gss_inv:**

19: em_gss_pln:**

20: em_gss_vrt:**

21: em_gss_fun:**

22: em_gss_mam:**

23: em_gss_mus:**

24: em_gss_pro:**

25: em_gss_rod:**

26: em_gss_phg:**

27: em_gss_vrl:**

28: gb_gss1:**

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118.8	7.4	1201	9	AL565455
2	115	7.1	1201	9	AL565455
3	109.6	6.8	1200	13	BX437758
4	105.8	6.6	1200	13	BX415878

C	5	105.8	6.6	1200	13	BX415878
C	6	104.6	6.5	1101	29	CNS00EVL
	7	103	6.4	400	29	DR46120T
	8	102.8	6.4	1056	13	BX415058
	9	102	6.3	797	28	AQ868440
	10	101.6	6.3	490	29	BX237235
C	11	100.6	6.2	581	29	BX217227
C	12	100.6	6.2	707	29	BX153480
C	13	100	6.2	698	29	BX200889
C	14	100	6.2	1201	13	BX446511
C	15	99	6.1	605	29	BX162842
C	16	98.4	6.1	1101	29	CNS00EVL
C	17	98.2	6.1	652	29	BX217370
C	18	98.2	6.1	739	29	BZ202464
C	19	98	6.1	797	28	AQ868440
C	20	98	6.1	1200	13	BX437758
C	21	97.6	6.1	371	29	BX241244
C	22	97.4	6.0	488	29	BX145735
C	23	97.2	6.0	675	29	DR48P23T
C	24	96.8	6.0	558	29	DR39M22T
C	25	96.8	6.0	574	29	BX149536
C	26	96.8	6.0	605	29	BX162842
C	27	96.6	6.0	548	29	BX185514
C	28	96.4	6.0	772	29	BX205878
C	29	96.2	6.0	400	29	DR46120T
C	30	96.2	6.0	604	29	BX204248
C	31	96.2	6.0	781	29	CC134240
C	32	96.2	6.0	829	29	BX143241
C	33	96	6.0	697	29	CNS04707
C	34	95.8	5.9	499	29	BX165235
C	35	95.6	5.9	697	29	CNS04707
C	36	95.4	5.9	404	29	BX176026
C	37	95.4	5.9	478	29	DR1G4T
C	38	95.4	5.9	520	29	BX224356
C	39	95.4	5.9	656	29	DR36P11T
C	40	95.4	5.9	739	29	BZ202464
C	41	95.2	5.9	486	29	BX225691
C	42	95.2	5.9	631	28	AZ133337
C	43	95.2	5.9	658	29	BX184360
C	44	95	5.9	594	29	BX139459
C	45	95	5.9	648	29	DR61B17T

ALIGNMENTS

RESULT 1

AL565455/c

LOCUS

DEFINITION

AL565455 Homo sapiens FETAL BRAIN Homo sapiens CDNA clone

CS0DF005YO18 3-PRIME, mRNA sequence.

AL565455

ACCESSION

AL565455

VERSION

EST.

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL565455 1201 bp mRNA linear EST 12-MAY-2003

AL565455 Homo sapiens FETAL BRAIN Homo sapiens CDNA clone

CS0DF005YO18 3-PRIME, mRNA sequence.

AL565455

ACCESSION

AL565455

VERSION

EST.

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Feb 16, 2001 this sequence version replaced gi:12916848.

Contact: Genoscope

Genoscope - Centre National de Sequenage

BP 191 91006 Evry cedex - France

Email: seqef@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 9232.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DF005H09NPL&cluster=9232.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

[illegible][illegible]

```
Qy 282 CTAGATAAACAGGCGCATTCAGATTCCTCTTGGTGCAGCTTGGAGAACCCATCTG 341
Db 658 AWATATATATATATATATATATATATATATATATATATATATATATATATATAT 717
Qy 342 GGCTTGGAGAGATTTACTTCTTGTTGATCTCTAGAGTACAGCTCCCTTAAGGCTGAGTC 401
Db 718 AWATATATATATATATATATATATATATATATATATATATATATATATATATAT 777
Qy 402 TAGTTTTTTTTTCTCCTACCAAAAAAAGTGCATATAAATAGCTTTATACAT 461
Db 778 AWATATATATATATATATATATATATATATATATATATATATATATATATATAT 837
Qy 462 ATAACTTTAATAAAAAATTCATCCCTTAAACACATAGTACAGAAATTTTCATAAA 521
Db 838 AWATATATATATATATATATATATATATATATATATATATATATATATATATAT 896
Qy 522 AAAATATGTTTATAATTTACATCCGTTACGGTAAAAAATGGATAAATGGGTATGGAG 581
Db 897 TAWATATATATATATATATATATATATATATATATATATATATATATATATATAT 956
Qy 582 TACTAGTAATTAAGGTTTCATTTGTTTAAAAAAGTAAAAAATTAATTTCTCTCTGATT 641
Db 957 TAWATATATATATATATATATATATATATATATATATATATATATATATATATAT 1016
Qy 642 TATATGAATGACATTTTTTTTGGACATGAGGCTATTGATTTTACCACCTTTTACACC 701
Db 1017 TAWATATATATATATATATATATATATATATATATATATATATATATATATATAT 1076
Qy 702 TTTCAAGCCATTCAGAGTGAATATAGATTTTGGGGGATCAACACAGAAATCATATAC 761
Db 1077 TAW-ATATATATATATATATATATATATATATATATATATATATATATATATATAT 1135
Qy 762 GATAACATGCTTATACATA 780
Db 1136 ATAWATATATATATATATATATATATATATATATATATATATATATATATATAT 1154

RESULT 3
BX437758/c
LOCUS BX437758 1200 bp mRNA linear EST 15-MAY-2003
DEFINITION BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008YB01
5-PRIME, mRNA sequence.
ACCESSION BX437758
VERSION BX437758.1
KEYWORDS GI:30773605
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li.W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequenage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSOCAP008CA01QPL1.
Location/Qualifiers
1. 1200
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOCAP008YB01"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
```

```
BASE COUNT 515 a 30 c 71 g 310 t 274 others
ORIGIN
Query Match 6.8%; Score 109.6; DB 13; Length 1200;
Best Local Similarity 32.2%; Pred.No. 0.04;
Matches 308; Conservative 183; Mismatches 462; Indels 3; Gaps 2;
Qy 1 TTTTAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1051
Db 1110 TTTTAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1051
Qy 61 TTAATAATATATATATATATATATATATATATATATATATATATATATATATATATAT 120
Db 1050 YTTTAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 991
Qy 121 TTAATAATATATATATATATATATATATATATATATATATATATATATATATATATAT 180
Db 990 AAAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 931
Qy 181 TTACGTTGATGCTGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 240
Db 930 TTAATAATATATATATATATATATATATATATATATATATATATATATATATATATAT 871
Qy 241 ACAATCCAAACCCCTTGGAGGCTTAATGTGTGTCGCGGATTTTCTAGATAAACAAGYGCCA 300
Db 870 TMTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 811
Qy 301 TTCAGGATTCCTTGGTGGAGCTTGGAGAACCCCTATCCTGGGCTTGGAGATTTTACATTC 360
Db 810 TTTTAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 751
Qy 361 TTGTTGATGCTCTAGATACAGCTCCTTAAGGCTGATCTAGTCTAGTCTAGTCTAGTCTAG 420
Db 750 WWTAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 692
Qy 421 TCCTACCAAAAAAAGTCAATAATATATATATATATATATATATATATATATATATATATAT 480
Db 691 MYTTCCTTWTAAWAAAWMTATMTATYTCYHWWAAACAAWAAWAAWAAWAAWAAWAAWAA 632
Qy 481 AAAAATTTATCCCTTAAACACATAGTAGAAATTTTCATATAAATAAATAAATAAATAA 540
Db 631 AAAMWHTCTTAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 572
Qy 541 ACATGCCGTACGGTAAATAAATGGATTAATTTGGGTATGGGTATGGGTATGGGTATGGGTAT 600
Db 571 AAATAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 512
Qy 601 TCATTGGTTTAAAAAACCTTAAAAAATAATTTCTCTCTCTGATTTATATATATATATATAT 660
Db 511 TCYTCYCTTWWAAWTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 452
Qy 661 TTGGAACATGAAGGCTATGATTTTACACCTTTTACACCTTTTACACCTTTTACACCTTTT 720
Db 451 YMMMAAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAA 394
Qy 721 TGAATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 780
Db 393 TTTCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 334
Qy 781 CCCGCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
Db 333 AAAMHTTCTTAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 274
Qy 841 ATAAACATCAAGCTCTAGCAAAATGTTGTTTGGCATGACACATTTTCATATATATATAT 900
Db 273 TTTTCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 214
Qy 901 AAGGATGCTATGATTCGATTACAAAAAATACTAATAATCTAGCAAAAGTTT 956
Db 213 TTTTCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 158
```

[illegible]


```

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0CAP004AD10NP1.
Location/Qualifiers
1. .1056
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP004Yig9"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT-6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT      454 a   53 c   56 g   318 t   175 others
ORIGIN
Query Match      6.4%;   Score 102.8;   DB 13;   Length 1056;

```

	Query match	96.4%; Score 102.8; DB 137; Length 1036;
	Best Local Similarity	35.8%; Pred. No. 0.19;
Matches 216; Conservative	99; Mismatches 289; Indels	0; Gaps 0;
QY	80 TATATATATATATATATATAATATTTTATTACCAATTTAAATATATATATATATA 139	
Db	360 TTTATATATATATATTTTGTATATATATATAWAAAAAATTTTTAAAAAAAAAAWTTKTA 419	
QY	140 TATATATATATATATATATAATATTTTATTATATTTATTCGTTGCTAAAAAAA 199	
	: : : :	
Db	420 TATAWTTTKTTTTTTWTATTTATAAAAAARWATATATTTTAAAAAATTTTTWWAAA 479	
QY	200 ATAAATATAATTTGGTTACCATTAAAAAGTCAPAAATATAGTACAATCCACCCCTTGAGA 259	
	: : : : : : : : : :	
Db	480 AAAAAATTTTTTATWAAATTTAAAAAWAAWAATTTWDKKKKKTATATAAAAAANAAAAWAA 539	

	260	GGTAAATGCTGTGGCGGATTTTCTAGATAAACCAAGYGCCATTACGATTTCTTCTGGTG	319
QY		: : :	:
	540	AAAAAAWTAATATATATWTATAKATAAIAAAAAAAAAAWAATATATAAAAAAAAAAAWA	599
Db		: : :	:
	320	CAGCTTGGAGAACCCTACTCCTGGCGCTTGAAGATTTACITCTTGTTGATGCTTCTAGAGT	379
QY		: : : :	:
	600	AWTTTTGWAAWATGTTTTTKTWTWTWTWAAWATATTTTWTATADATRTKTTTWTAT	659
Db		: : :	:
	380	ACAGCTCCCTAAGGCCTGTAGTCAGTTTTTTTTCATCCTCCTACCAAAAAAAAAAAAA	439
QY		: : : :	:
	660	ADWDATWAAWATATWTWTRTGRKGAWIAAAAANKTAKKTATWAAAAIAAAAAAAAAA	719
Db		: : :	:
	440	GTCATAATATAGTTTATACATATAACTTTAAATAAAAAATAAAAAATTTTCATCCCCAAA	499
QY		: : :	:
	720	AAAAAAAAAAAAAAAAAAVATTTTTTTTTTWAAAAAAAAAAATNNNNNNNTTAAWATWWMTT	779
Db		: : :	:

500	ACATAGTAGAGAAATTTTCATAAAAAAATAATATGTTTATAATTTACATGCCGTTTACGGTAAAA	359
780	TTTTTTTTTTTTWWWDGDDDDDDTT	839
560	AATGGATAAATGGGTATCGGTACTAGTAAATAATAGGTTTCATGGTTTAAAAAACATA	619
	:: :: :: :: ::	
840	WWWWWWWWAAAAAATAATTTTAAATTTTATAAWWAAAAATGTTTTTAAWAAAAA	899
620	AAAAATAATTTCTCTCCTGATTATATGAATAATGACATTTTTTTTGAACAATGAAGGTTAT	679
900	AAAAAAAWTTTAAWAAAWMTTWAAAWMTWKNAAAAAAATAAAAAAATAAAAAATTTTT	959
680	GATT	683
960	TATT	963
RESULT 9		

LOCUS AQ868440
DEFINITION nbeb027F06r CUGI Rice BAC Library (ECORI)
FEATURES: nbeb027F06r genomic clone nbeb027F06r, genomic survey cultivar-group) genomic clone nbeb027F06r, genomic survey sequence.
linear GSS 03-NOV-1999

ACCESSION AQ868440
 VERSION AQ868440.1 GI:6218891
 KEYWORDS GSS.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 797)
 REFERENCE Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Rice Genome
 UNPUBLISHED
 CONTACT: Wing RA
 CLEMSON UNIVERSITY GENOMICS INSTITUTE
 100 JORDAN HALL, CLEMSON, SC 29634, USA
 TEL: 864 656 7288
 FAX: 864 656 4293
 EMAIL: rwing@clcmson.edu
 SEQ PRIMER: GGAACAGCTATGACCATG
 CLASS: BAC ends
 HIGH QUALITY SEQUENCE START: 41
 HIGH QUALITY SEQUENCE STOP: 329.
 FEATURES
 source
 1..797
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="nbeb0027F06r"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice BAC Library (ECORI)"
 /note="vector: pBACIndigo; Site 1: EcoRI; Site 2: EcoRI;
 Rice is the most important food crop in the world. Half of
 the world population, especially those inhabiting highly
 populated areas of the humid tropics and subtropics, rely
 on rice as their primary source of carbohydrate.
 Monocytodendous rice is a diploid plant (2n=24) with a
 haploid genome equivalent of 431 Mbp (Arumuganathan and
 Earle, 1991). The relatively small genome of rice, three
 times larger than that of Arabidopsis, makes it suitable
 for genomic studies. In order to facilitate positional
 cloning, physical mapping and genome sequencing of rice,
 we have constructed a BAC library from Oryza sativa,
 Nipponbare variety using EcoRI as the cloning enzyme. The
 library contains 55,296 clones with an average insert size
 of 121 kb providing approximately 15 haploid genome
 equivalents. The deep coverage allows the isolation a
 particular sequence with a probability of 99.9 %. Three
 high density filters, each containing 18,432 clones
 (doubly spotted), represent the whole library for colony
 screening and can be requested from the Clemson University
 BAC/EST Resource Center (www.genome.clemson.edu)."
 BASE COUNT 285 a 179 c 48 g 275 t 10 others
 ORIGIN
 Query Match 6.3%; Score 102; DB 28; Length 797;
 Best Local Similarity 59.8%; Pred. No. 0.26; Indels 0; Gaps 0;
 Matches 171; Conservative 0; Mismatches 115;
 Oy 2 TTTATAATATTTAAAGCTGATATAATTTTGGCATCTATATATAAGCCCACTACCAATT 61
 Db 220 TGTAATATATTTTGTCATATATATATATCTATATACATATGCTTGCATACATATA 279
 Oy 62 TAAAT 121
 Db 280 TATGTATAAATACATA 339
 Oy 122 TAAAT 181
 Db 340 TAT 399

Oy 182 TACCTTGATGGTAAAAATAAATAATATATTTTGTACCACTTTAAAGTCATAAATAGTA 241
 Db 400 ATATATATATATATATATATATATATATATATTTTAAATATAATATATACACACACA 459
 Oy 242 CAATCCCAACCCCTTGGAGAGGTTAATGTGTGCGGATTTTCTAGAT 287
 Db 460 CACACACACATATATATATATATATATATATATAAATTTCTATATAT 505
 RESULT 10
 LOCUS BX237235 490 bp DNA linear GSS 29-JAN-2003
 DEFINITION Danio rerio genomic clone DKEY-283H15, genomic survey sequence.
 ACCESSION BX237235
 VERSION BX237235.1 GI:28159569
 KEYWORDS GSS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 490)
 REFERENCE Humphray, S.J., Huckle, E. and Durham, J.L.
 Direct SubMISSION
 TITLE Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
 JOURNAL Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Unpublished
 COMMENT This sequence was generated from the SP6 end of BAC 283H15. 283H15
 is part of the Daniokey BAC Library created by R. Plasterk and N.V.
 Keygene. Further details:
 http://www.sanger.ac.uk/Projects/D_rerio/
 FEATURES
 source
 1..490
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="DKEY-283H15"
 /tissue_type="Testis"
 /note="vector pIndigoBAC-536"
 BASE COUNT 204 a 32 c 35 g 219 t
 ORIGIN
 Query Match 6.3%; Score 101.6; DB 29; Length 490;
 Best Local Similarity 72.8%; Pred. No. 0.36;
 Matches 131; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 Oy 59 ATTTAAAT 118
 Db 220 AT 279
 Oy 119 ATTTAAAT 178
 Db 280 AT 339
 Oy 179 TATTACCGTGTGCTAAATAAATAAATAATATATTTGTACCACTTTAAAGTCATAAATA 238
 Db 340 TATTTTA 399
 RESULT 11
 LOCUS BX217227/c 581 bp DNA linear GSS 29-JAN-2003
 DEFINITION Danio rerio genomic clone DKEY-258I22, genomic survey sequence.
 ACCESSION BX217227
 VERSION BX217227.1 GI:28049113
 KEYWORDS GSS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 581)

THIS PAGE BLANK (USPTO)